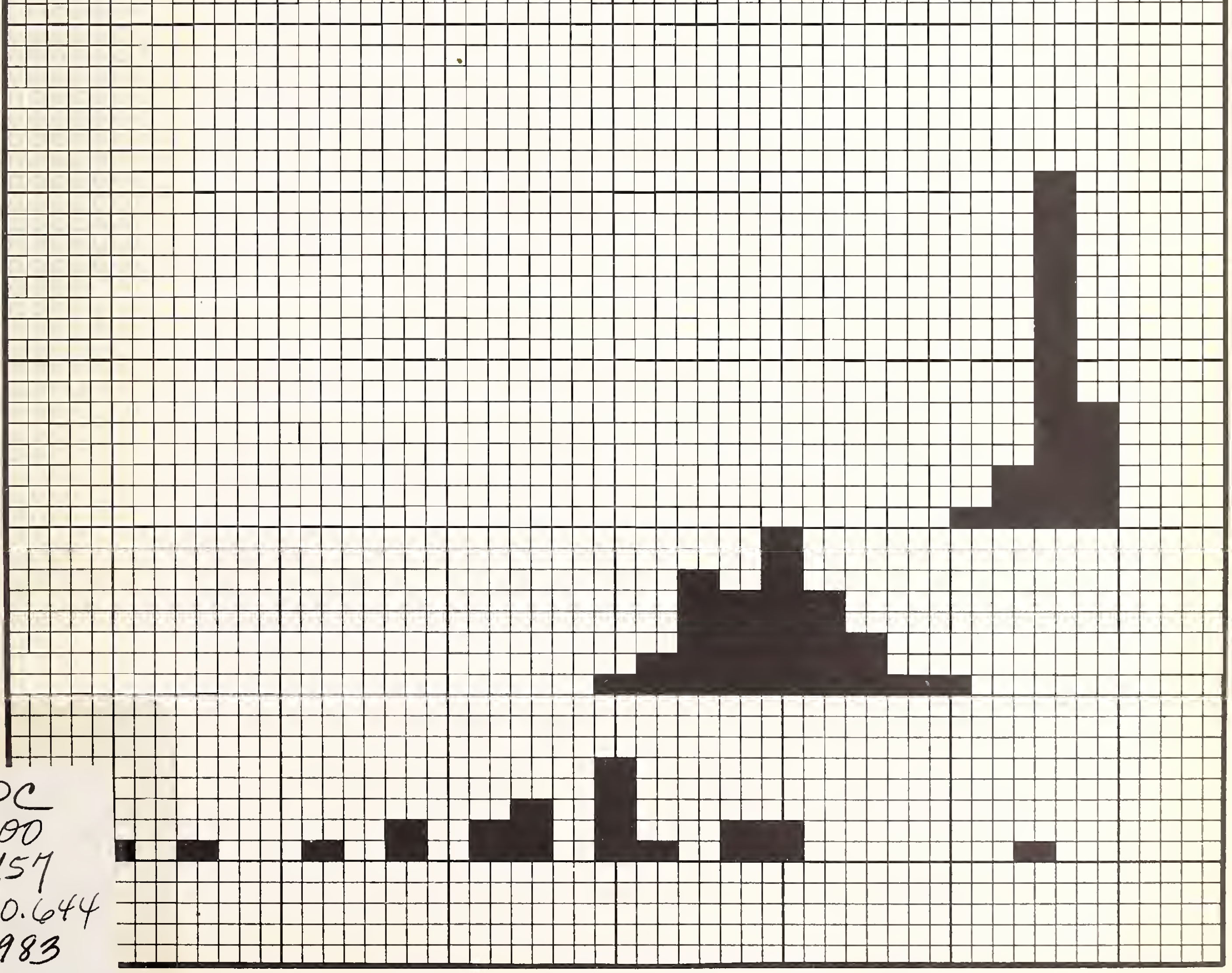
NAT'L INST. OF STAND & TECH
A11105 966895

NIST
PUBLICATIONS

NBS SPECIAL PUBLICATION 644

U.S. DEPARTMENT OF COMMERCE/National Bureau of Standards

Expression of the Uncertainties of Final Measurement Results: Reprints



QC
100
.U57
no.644
1983

NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, and the Institute for Computer Sciences and Technology.

THE NATIONAL MEASUREMENT LABORATORY provides the national system of physical and chemical and materials measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; conducts materials research leading to improved methods of measurement, standards, and data on the properties of materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

Absolute Physical Quantities[2] — Radiation Research — Chemical Physics — Analytical Chemistry — Materials Science

THE NATIONAL ENGINEERING LABORATORY provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

Applied Mathematics — Electronics and Electrical Engineering[2] — Manufacturing Engineering — Building Technology — Fire Research — Chemical Engineering[2]

THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

Programming Science and Technology — Computer Systems Engineering.

[1]Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted;
mailing address Washington, DC 20234.
[2]Some divisions within the center are located at Boulder, CO 80303.

Expression of the Uncertainties of Final Measurement Results: Reprints

Churchill Eisenhart

Harry H. Ku

R. Collé

National Bureau of Standards
Washington, DC 20234

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, Secretary

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Director

Issued January 1983

Library of Congress Catalog Card Number: 82-600655

National Bureau of Standards Special Publication 644
Natl. Bur. Stand. (U.S.), Spec. Publ. 644, 19 pages (Jan. 1983)
CODEN: XNBSAV

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1983

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402
Price $3.00
(Add 25 percent for other than U.S. mailing)

Foreword

The reporting of final measurement results, and the uncertainties associated with the measurement processes used to obtain these results, has always been and continues to be a source of difficulty. The three articles reprinted in this publication are collected here as a convenient reference source for experimenters who must face the difficult task of deciding how to express measurement uncertainties. The philosophical basis, general guidelines, and specific recommendations for expressing uncertainties contained within these articles have evolved at NBS over a period of many years.

The first article originally appeared in *Science* in 1968. This article develops the underlying basis and general guidelines on the forms of expression needed for uncertainty statements, and presents specific recommendations for four distinct cases: (i) when both systematic error and imprecision are negligible; (ii) when systematic error is not negligible, and imprecision is negligible; (iii) when neither systematic error nor imprecision is negligible; and (iv) when systematic error is negligible, and imprecision is not negligible.

The second article, written as a companion to the first, originally appeared in a 1968 issue of *M&D: Measurements and Data.* It gives a condensed summary of the recommendations presented in the first article, and provides tabular guides to commonly used statements of imprecision, systematic error, and uncertainty.

The third article is a postscript to the two preceding articles, and was prepared in 1980 for an internal NBS communications manual. It reinforces the major thrust and content of the earlier articles, but includes more recent thought particularly in regard to overall uncertainty statements.

The first two articles have since been reprinted in several NBS publications including Special Publication 300, Volume 1, *Precision Measurement and Calibration: Statistical Concepts and Procedures* (Harry H. Ku, ed., 1969). The 1980 NBS communications manual incorporated the second and third articles, but did not reprint the first article. Furthermore, this manual is not accessible outside NBS. This special publication, therefore, collects all three articles, for the first time, in one convenient source which is available to the many scientists and engineers throughout the entire measurement community.

Abstract

This publication reprints and collects in one convenient source three articles, by NBS authors, that present a philosophical basis, general guidelines, and specific recommendations for expressing the uncertainties of final measurement results.

Key words: accuracy; errors; measurement uncertainty; precision; reporting of measurement data; systematic error; uncertainties.

Contents

Expression of the Uncertainties of Final Results

Clear statements of the uncertainties of reported values are needed for their critical evaluation.

Churchill Eisenhart

Measurement of some property of a thing in practice always takes the form of a sequence of steps or operations that yield as an end result a number that serves to represent the amount or quantity of some particular property of a thing—a number that indicates how much of this property the thing has, for someone to use for a specific purpose. The end result may be the outcome of a single reading of an instrument, with or without corrections for departures from prescribed conditions. More often it is some kind of average, for example, the arithmetic mean of a number of independent determinations of the same magnitude, or the final result of a least squares "reduction" of measurements of a number of different magnitudes that bear known relations with one another in accordance with a definite experimental plan. In general, the purpose for which the answer is needed determines the precision or accuracy required and ordinarily also the method of measurement employed.

Although the accuracy required of a reported value depends primarily on the *intended* use, or uses, of the value, one should not ignore the requirements of other uses to which it is likely to be put. A reported value whose accuracy is entirely unknown is worthless.

Strictly speaking, the actual *error* of a reported value, that is the magnitude and sign of its deviation from the truth (*1*), is usually unknowable. Limits to this error, however, can usually be inferred—with some risk of being incorrect—from the precision of the measurement process by which the reported value was obtained, and from reasonable limits to the possible bias of the measurement process. The *bias*, or *systematic error*, of a measurement process is the magnitude and direction of its tendency to measure something other than what was intended; its *precision* refers to the typical closeness together of successive independent measurements of a single magnitude generated by repeated applications of the process under specified conditions; and its *accuracy* is determined by the closeness to the true value characteristic of such measurements.

Precision and accuracy are inherent characteristics of the measurement process employed and not of the particular end result obtained. From experience with a particular measurement process and knowledge of its sensitivity to uncontrolled factors, one can often place reasonable bounds on its likely systematic error (bias). It is also necessary to know how well the particular value in hand is likely to agree with other values that the same measurement process might have provided in this instance, or might yield on remeasurement of the same magnitude on another occasion. Such information is provided by the estimated *standard error* (*2*) of the reported value, which measures (or is an index of) the characteristic disagreement of repeated determinations of the same quantity by the same method, and thus serves to indicate the precision (strictly, the imprecision) of the reported value (*3*).

Four Distinct Forms of Expression Needed

The uncertainty of a reported value is indicated by stating credible limits to its likely inaccuracy. No single form of expression for these limits is universally satisfactory. In fact, different forms of expression are recommended, which will depend on the relative magnitudes of the imprecision and likely bias, and their relative importance in relation to the intended use of the reported value, as well as to other possible uses to which it may be put (*4*).

Four distinct cases need to be recognized: (i) both systematic error and imprecision negligible, in relation to the requirements of the intended and likely uses of the result; (ii) systematic error not negligible, imprecision negligible; (iii) neither systematic error nor imprecision negligible; and (iv) systematic error negligible, imprecision not negligible.

Specific recommendations with respect to each of these cases are made below. General guidelines upon which these specific recommendations are based are discussed in the following paragraphs.

Perils of Shorthand Expressions

Final results and their respective uncertainties should be reported in sentence form whenever possible. The shorthand form "$a \pm b$" should be avoided in abstracts and summaries; and never used without explicit explanation of its connotation. If no explanation is given, many persons will take $\pm b$ to signify bounds to the inaccuracy of a. Others may assume that b is the "standard error," or the "probable error," of a, and hence the uncertainty of a is at least $\pm 3b$, or $\pm 4b$, respectively. Still others may take b to be an indication merely of the imprecision of the individual measurements, that is, to be the "standard deviation," or the "average deviation," or the "probable error" of a single observation. Each of these interpretations reflects a practice of which instances can be found in current scientific literature. As a step in the direction of reducing this current confusion, it is recommended that the use of "$a \pm b$" in presenting results be limited to that sanctioned for the case of tabular results in the fourth recommendation of the section below headed "Systematic error not negligible, imprecision negligible."

The author is a senior research fellow and former chief of the Statistical Engineering Laboratory at the National Bureau of Standards, Washington, D.C. 20234. The recommendations presented in this paper have evolved at the Bureau over a period of many years and are made public here for general information, and to educe comments and suggestions.

Imprecision and Systematic Error Require Separate Treatment

Since imprecision and systematic error are distinctly different components of inaccuracy, and are subject to different treatments and interpretations in usage, two numerics respectively expressing the imprecision and bounds to the systematic error of the reported result should be used whenever both of these errors are factors requiring consideration. Such instances are discussed in the section below for the case of "Neither systematic error nor imprecision negligible."

In quoting a reported value and its associated uncertainty from the literature, the interpretation of the uncertainty quoted should be stated if given by the author. If the interpretation is not known, a remark to this effect is in order. This practice may induce authors to use more explicit formulations of their statements of uncertainty.

Standard Deviation and Standard Error

The terms *standard deviation* and *standard error* should be reserved to denote the canonical values for the measurement process, based on considerable recent experience with the measurement process or processes involved. When there is insufficient recent experience, an estimate of the standard error (standard deviation) must of necessity be computed by recognized statistical procedures from the same measurements as the reported value itself. To avoid possible misunderstanding, in such cases, the term "computed (or estimated) standard error" ("computed standard deviation") should be used. A formula for calculating this computed standard error is given in the section below for the case of "Neither systematic error nor imprecision negligible."

Uncertainties of Accepted Values of Fundamental Constants or Primary Standards

If the uncertainty in the accepted value of a national primary standard or of some fundamental constant of nature (for example, in the volt as maintained at the National Bureau of Standards, or in the acceleration of gravity *g* on the Potsdam basis) is an important source of systematic error affecting the measurement process, no allowance for possible systematic error from this source should be included ordinarily in evaluating overall bounds to the systematic error of the measurement process. Since the error concerned, whatever it is, affects all results obtained by the method of measurement involved, to include an allowance for this error would be to make everybody's results appear unduly inaccurate relative to each other. In such instances one should state: (i) that measurements obtained by the process concerned are expressed in terms of the volt (or the kilogram, or other unit) "as maintained at the National Bureau of Standards," or (ii) that the indicated bounds to the systematic error of the process are exclusive of the uncertainty of the stated value adopted for some particular constant or quantity. An example of the latter form of statement is:

. . . neglecting the uncertainty of the value 6.6256×10^{-34} joule seconds adopted for Planck's constant.

Systematic Error and Imprecision Both Negligible

In this case the reported result should be given, after rounding, to the number of significant figures consistent with the accuracy requirements of the situation, together with an explicit statement of its accuracy. An example is:

. . . the wavelengths of the principal visible lines of mercury-198 have been measured relative to the 6057.802106 Å (angstrom units) line of krypton-98, and their values in vacuum are

5792.2685 Å
5771.1984 Å
5462.2706 Å
4359.5625 Å
4047.7146 Å

correct to eight significant figures.

It needs to be emphasized that if no statement of accuracy or precision accompanies a reported number, then, in accordance with the usual conventions governing rounding, this number will ordinarily be interpreted as being accurate within ±½ unit in the last significant figure given; that is, it will be understood that its inaccuracy before rounding was less than ± 5 units in the next place. The statement "correct to eight significant figures" is included explicitly in the foregoing example, rather than left to be understood in order to forestall any concern that an explicit statement of lesser accuracy was inadvertently omitted.

Systematic Error Not Negligible, Imprecision Negligible

When the imprecision of a result is negligible, but the inherent systematic error of the measurement process concerned is not negligible, then the following rules are recommended:

1) Qualification of a reported result should be limited to a single quasi-absolute type of statement that places bounds on its inaccuracy.

2) These bounds should be stated to no more than two significant figures.

3) The reported result itself should be given (that is, rounded) to the last place affected by the stated bounds (unless it is desired to indicate and preserve such relative accuracy or precision of a higher order that it may possess for certain particular uses).

4) Accuracy statements should be given in sentence form in all cases, except when a number of results of different accuracies are presented, for example, in tabular arrangement. If it is necessary or desirable to indicate the respective accuracies of a number of results, the results should be given in the form $a \pm b$ (or $a\ ^{+b}_{-c}$, if necessary) with an appropriate explanatory remark (as a footnote to the table, or incorporated in the accompanying text) to the effect that the $\pm b$, or $^{+b}_{-c}$, signify bounds to the systematic errors to which the a's may be subject.

5) The fact that the imprecision is negligible should be stated explicitly.

The particular form of the quasi-absolute type of statement employed in a given instance will depend ordinarily on personal taste, experience, current and past practice in the field of activity concerned, and so forth. Some examples of good practice are:

. . . is (are) not in error by more than 1 part in (x).
. . . is (are) accurate within ± (x units) [or ± (x) percent].
. . . is (are) believed accurate within (. . . . .).

Positive wording, as in the first two of these quasi-absolute statements, is appropriate only when the stated bounds to the possible inaccuracy of the reported value are themselves reliably established. However, when the indicated bounds are somewhat conjectural, it is desirable to signify this fact (and put the reader on guard) by inclusion of some modifying expression such as "believed," "considered," "estimated to be," "thought to be," and

so forth, as exemplified by the third of the foregoing examples.

The term *uncertainty* may sometimes be used effectively to achieve a conciseness of expression otherwise difficult or impossible to attain. Thus, one might make a statement such as:

> The uncertainties in the above values are not more than $\pm$ 0.5°C in the range 0°C to 1100°C, and then increase to $\pm$ 2°C at 1450°C,

or

> The uncertainty in this value does not exceed . . . excluding (or, including) the uncertainty of . . . in the value . . . adopted for the (reference standard involved).

A statement giving numerical limits of uncertainty as in the above should be followed by a brief discussion telling how the limits were derived.

Finally, the following forms of quasi-absolute statements are considered poor practice, and are to be avoided:

> The accuracy of . . . is 5 percent.
> The accuracy of . . . is $\pm$ 2 percent.

These are presumably intended to mean that the result concerned is not inaccurate, that is, not in error, by more than 5 percent or 2 percent, respectively, but they explicitly state the opposite.

Neither Systematic Error Nor Imprecision Negligible

When neither the imprecision nor the systematic error of a result are negligible, then the following rules are recommended:

1) A reported result should be qualified by a quasi-absolute type of statement that places bounds on its systematic error, and a separate statement of its standard error or its probable error, or of an upper bound thereto, whenever a reliable determination of such value or bound is available. Otherwise a computed value of the standard error, or, probable error, so designated, should be given together with a statement of the number of degrees of freedom on which it is based.

2) The bounds to its systematic error and the measure of its imprecision should be stated to no more than two significant figures.

3) The reported result itself should be stated at most to the last place affected by the finer of the two qualifying statements (unless it is desired to indicate and preserve such relative accuracy or precision of a higher order that it may possess for certain particular uses).

4) The qualification of a reported result with respect to its imprecision and systematic error should be given in sentence form, except when results of different precision or with different bounds to their systematic errors are presented in tabular arrangement. If it is necessary or desirable to indicate their respective imprecisions or bounds to their respective systematic errors, such information may be given in a parallel column or columns, with appropriate identification.

Here, and in the next section, the term *standard error* is to be understood as signifying the standard deviation of the reported value itself, not as signifying the standard deviation of the single determination (unless, of course, the reported value is simply the result of a single determination).

The above recommendations should not be construed to exclude the presentation of a quasi-abolute type of statement placing bounds on the inaccuracy, that is, on the overall uncertainty, of a reported value, provided that separate statements of its imprecision and its possible systematic error are included also. To be in good taste, the bounds indicating the overall uncertainty should not be numerically less than the corresponding bounds placed on the systematic error outwardly increased by at least three times the standard error. The fourth of the following examples of good practice is an instance at point:

> The standard errors of these values do not exceed 0.000004 inch, and their systematic errors are not in excess of 0.00002 inch.

> The standard errors of these values are less than (x units), and their systematic errors are thought to be less than $\pm$ (y units). No additional uncertainty is assigned for the conversion to the chemical scale since the adopted conversion factor is taken as 1.000275 exactly.

> . . . with a standard error of (x units), and a systematic error of not more than $\pm$ (y units).

> . . . with an overall uncertainty of $\pm$ 3 percent based on a standard error of 0.5 percent and an allowance of $\pm$ 1.5 percent for systematic error.

When a reliably established value for the relevant standard error is available, and the dispersion of the present measurements is in keeping with this experience, then this canonical value of the standard error should be used (*5*). If such experience indicates that the standard error is subject to fluctuations greater than the intrinsic variation of such a measure, then an appropriate upper bound should be given, for example, as in the first two of the above examples, or by changing "a standard error . . ." in the third and fourth examples to "an upper bound to the standard error . . ."

When there is insufficient recent experience with the measurement processes involved, an estimate of the standard error must of necessity be computed by recognized statistical procedures from the same measurements as the reported value itself. It is essential that such computations be carried out according to an agreed-upon standard procedure, and the results thereof presented in sufficient detail to enable the reader to form his own judgment, and make his own allowances for their inherent uncertainties. To avoid possible misunderstanding, in such cases, first, the term *computed standard error* should be used; second, the estimate of the standard error employed should be that obtained from

$$\text{estimate of standard error} = \left(\frac{\text{sum of squared residuals}}{n\nu}\right)^{1/2}$$

where n is the (effective) number of completely independent determinations of which a is the arithmetic mean (or other appropriate least-squares adjusted value) and ν is the number of degrees of freedom involved in the sum of squared residuals (that is, the number of residuals minus the number of fitted constants or other independent constraints on the residuals); and third, the number of degrees of freedom should be explicitly stated. If the reported value a is the arithmetic mean, then:

$$\text{estimate of standard error} = (s^2/n)^{1/2}$$

where

$$s^2 = \sum_{i=1}^{n} (x_i - a)^2/(n-1)$$

and n is the number of completely independent determinations of which a is the arithmetic mean. For example:

> . . . which is the arithmetic mean of (n) independent determinations and has a standard error of . . .

> . . . with an overall uncertainty of $\pm$ 5.2 km/sec based on a standard error of 1.5 km/sec and estimated bounds of $\pm$ 0.7 km/sec on the systematic error. (The figure 5.2 is equal to 0.7 plus 3 times 1.5.)

or, if based on a computed standard error,

> The computed probable error (or, standard error) of these values is (x units),

based on (ν) degrees of freedom, and the systematic error is estimated to be less than $\pm$ (y units).

. . . with an overall uncertainty of $\pm$ 7 km/sec derived from bounds of $\pm$ 0.7 km/sec on the systematic error and a computed standard error of 1.5 km/sec based on 9 degrees of freedom. [The number 7 is approximately equal to $0.7 + (4.3 \times 1.5)$, where 4.3 is the value of Student's t for 9 degrees of freedom exceeded in absolute value with 0.002 probability. As $\nu \rightarrow \infty$, $t_{.002}\ (\nu) \rightarrow 3.090$.]

When the reported value is the result of a complex measurement process and is obtained as a function of several quantities whose standard errors have been computed, these several quantities and their standard errors should usually be reported, together with a description of the method of computation by which the standard errors were combined to provide an overall estimate of imprecision for the reported value.

Systematic Error Negligible, Imprecision Not Negligible

When the systematic error of a result is negligible but its imprecision is not, the following rules are recommended:

1) Qualification of a reported value should be limited to a statement of its standard error or of an upper bound thereto, whenever a reliable determination of such value or bound is available. Otherwise a computed value of the standard error, so designated, should be given together with a statement of the number of degrees of freedom on which it is based.

2) The standard error or upper bound thereto, should be stated to not more than two significant figures.

3) The reported result itself should be stated at most to the last place affected by the stated value or bound to its imprecision (unless it is desired to indicate and preserve such relative precision of a higher order that it may possess for certain particular uses).

4) The qualification of a reported result with respect to its imprecision should be given in sentence form, except when results of different precision are presented in tabular arrangement and it is necessary or desirable to indicate their respective imprecisions in which event such information may be given in a parallel column or columns, with appropriate identification.

5) The fact that the systematic error is negligible should be stated explicitly.

The above recommendations should not be construed to exclude the presentation of a quasi-absolute type of statement placing bounds on its possible inaccuracy, provided that a separate statement of its imprecision is included also. To be in good taste, such bounds to its inaccuracy should be numerically equal to at least three times the stated standard error. The fourth of the following examples of good practice is an instance at point.

The standard errors of these values are less than (x units).

. . . with a standard error of (x units).

. . . with a computed standard error of (x units) based on (ν) degrees of freedom.

. . . with an overall uncertainty of $\pm$ 4.5 km/sec derived from a standard error of 1.5 km/sec. (The figure 4.5 is equal to 3×1.5.)

or, if based on a computed standard error,

. . . with an overall uncertainty of $\pm$ 6.5 km/sec derived from a computed standard error of 1.5 km/sec (based on 9 degrees of freedom). (The number 6.5 is equal to 4.3×1.5, where 4.3 is the value of Student's t for 9 degrees of freedom exceeded in absolute value with 0.002 probability. As $\nu \rightarrow \infty$, $t_{.002}\ (\nu) \rightarrow 3.090$.)

The remarks with regard to a computed standard error in the preceding section apply with equal force to the last two examples above.

Conclusion

The foregoing recommendations call for fuller and sharper detail than is general in common pactice. They should be regarded as minimum standards of good practice. Of course, many instances require fuller treatment than that recommended here.

Thus, in the case of determinations of the "fundamental physical constants" and other basic properties of nature, the author or authors should give a detailed account of the various components of imprecision and systematic error, and list their respective individual magnitudes in tabular form, so that (i) the state of the art will be more clearly revealed, (ii) each individual user of the final result may decide for himself which of the indicated components of imprecision or systematic error are, or are not, relevant to his use of the final result, and (iii)—most important—the final result itself or its uncertainty can be modified appropriately in the light of later advances. This is, and has long been, the practice followed in the best reports of fundamental studies, but current efforts to prepare critically evaluated standard reference data have revealed that far too great a fraction of the data in the scientific literature "cannot be critically evaluated because the minimum of essential information is not present" (*6*).

References and Notes

1. The true value defined conceptually by an exemplar measurement process, or the target value intended in a practical measurement process.
2. The standard error is the standard deviation of the probability distribution of estimates (that is, reported values) of the quantity that is being measured. See M. G. Kendall and W. R. Buckland, *A Dictionary of Statistical Terms* (Hafner, New York, 1957).
3. For a comprehensive discussion on precision and accuracy, and a selected bibliography of 80 references, see C. Eisenhart, "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems," *J. Res. Nat. Bur. Std.* **67C**, No. 2, 161–187 (1963). (Reprints are available upon request.)
4. The essential elements of the present recommendations first appeared in a 1955 National Bureau of Standards task group report prepared principally by Malcolm W. Jensen (Office of Weights and Measures), Leroy W. Tilton (Optics and Metrology Division), and Churchill Eisenhart (Applied Mathematics Division), which was based for the most part on detailed recommendations developed some years earlier by Dr. Tilton for the internal guidance of the Optics and Metrology Division. In September 1961, new introductory material was added to the recommendations of the 1955 task group; a few minor changes were made in the illustrative examples, and the resulting revised version was circulated as a working paper of the Subcommittee on Accuracy Statements of the NBS Testing and Calibration Committee. This 1961 version was incorporated without essential change as chapter 23, "Expression of the Uncertainties of Final Results," of NBS Handbook 91, *Experimental Statistics* (U.S. Government Printing Office, Washington, 1963), reprinted with corrections in 1966. (This handbook brought together in a single volume the material on experimental statistics prepared at the National Bureau of Standards for the U.S. Army *Ordnance Engineering Design Handbook*, and printed in 1962 for limited distribution as U.S. Army Ordnance Corps Pamphlets ORDP 20-110 through 20-114. Subsequently, when these five pamphlets became parts of the *AMC Engineering Design Handbook*, they were designated Army Materiel Command Pamphlets AMCP 706-110 through 706-114.)

   In the present version, the content of chapter 23 has been rearranged and, in order to be more appropriate to calibration work, more explicit consideration has been given to the case where the value of the standard deviation σ of the measurement process involved has been well established by recent past experience. A terse summary of the principal recommendations of the present paper in the form of a text figure (Fig. 1) is contained in H. H. Ku, "Expressions of Imprecision, Systematic Error, and Uncertainty Associated with a Reported Value," to be published in *Measurements and Data*. The earlier versions were addressed primarily to the case of isolated experiments or tests, where the relevant value of σ is usually unknown in advance, and the statistical uncertainty of the final results must therefore be expressed entirely in terms of quantities derived from the data of the experiment itself.
5. The control chart is an invaluable tool in providing justification for the use of a canonical value of the standard error. See, for example, ASTM Manual on *Quality Control of Materials* (American Society for Testing and Materials, Philadephia, 1951).
6. L. M. Branscomb, "The misinformation explosion: Is the literature worth reviewing?," a talk presented to the Philosophical Society of Washington, 17 November 1967, and to be published in *Scientific Research*.

EXPRESSIONS OF IMPRECISION, SYSTEMATIC ERROR, AND UNCERTAINTY ASSOCIATED WITH A REPORTED VALUE

HARRY H. KU, National Bureau of Standards

The work of a calibration laboratory may be thought of as a sequence of operations that result in the collection, storage, and transmittal of information. In making a statement of uncertainty of the result of calibration, the calibration laboratory transmits information to its clients on the particular item calibrated.

It is logical, then, to require the transmitted information to be meaningful and unambiguous, and to contain all the relevant information in the possession of the laboratory. *The information content of the statement of uncertainty determines, to a large extent, the worth of the calibrated value.*

A common deficiency in many statements of uncertainty is that they do not convey all the information a calibration laboratory has to offer, information acquired through much ingenuity and hard work. This deficiency usually originates in two ways:

1. Loss of information through oversimplification, and

2. loss of information through the inability of the laboratory to take into account information accumulated from its past experience.

With the increasingly stringent demands for improved precision and accuracy of calibration work, calibration laboratories as a whole just cannot afford such luxury.

Traceability to the national standards, accuracy ratios, and class tolerance requirements are simplified concepts that aim to achieve different degrees of accuracy requirements. These concepts and the resulting statements are useful on certain occasions, but fail whenever the demand is exacting. The general practice of obliterating all the identifiable components of uncertainty, by combining them into an overall uncertainty, just for the sake of simplicity, is another case in point. After all, if the calibration laboratory reports all the pertinent information in separate components, the user can always combine them or use them individually, as he sees fit. On the other hand, if the user is given only one number, he can never disentangle this number into its various components. Since the information buried under these oversimplified statements is available, and may well be useful to sophisticated customers, such practices result in substantial waste of effort and resources.

In calibrating an item by repeating the same calibration procedure, the calibration laboratory gains increments of information about its calibration *system*. These increments of information are quantified and accumulated for the benefit of the calibration laboratory. If the precision of the calibration process remains unchanged, the statistical measure of dispersion (s) – i.e., the standard deviations computed from these sets of data – can be pooled together, weighted by their respective degrees of freedom. When many such increments of information are combined, an accepted or canonical value of standard deviation (σ) is established. This established (canonical) value of standard deviation characterizes the precision of the calibration process, and is treasured information in any calibration laboratory.

Hence, the canonical value of standard deviation is the quantification of information accumulated from past experiences of the calibration laboratory, and is an essential element of the statement of uncertainty. The standard deviation (s) computed from the current calibration is used to check the precision of current work, and to add to the pool of information on the process, but certainly does not represent all the information available in the possession of an established calibration laboratory. Only by passing its accumulated information to the users is the calibration laboratory performing a complete service.

STATEMENT OF UNCERTAINTY

In the preparation of a statement of uncertainty, it is helpful to bear in mind that:

1. The derivation of a statement of uncertainty has as its foundation the work done in the laboratory, and is based on information accumulated from past experience, and

2. In general, information is lost through oversimplification, and demands for improved precision and accuracy cannot be met with simplified statements of uncertainty.

Unless a statement of uncertainty is well formulated and supported, it is difficult to say what is meant by the statement, a difficulty frequently encountered. Since the evaluation of uncertainty is part and parcel of the high standard of work of a calibration laboratory, the statement of uncertainty deserves all the attention required to make the statement both realistic and useful. To this end, Tables 1, 2 and 3 give terms and expressions compiled as a ready reference for those who are searching for some appropriate format or wording, to carry out the thoughts expressed. They summarize the recommended practices on expression of uncertainties as given in Chapter 23 of NBS Handbook 91. A revised version of this chapter with the title "Expression of Uncertainties of Final Results" by Churchill Eisenhart may be found in NBS Special Publication 300-1. Figure 1 gives a condensed summary of this material. Tables 1, 2, and 3 give details of forms of imprecision, systematic error, and uncertainty statements.

TABLE 1 – IMPRECISION STATEMENTS

Value reported	Index or Measure of Error	Remarks
Precision of a measurement (calibration) process	(a). Standard deviation (σ) of a single determination (observation)	σ (or s with the associated degrees of freedom[1]) is of main interest as an index of precision of the measurement process. If the average of n such measurements is also reported, see (b) below.
Arithmetic mean ($\bar{x}_n$) of n numbers	(b). Standard error ($\sigma/\sqrt{n}$) of the reported value	$\bar{x}_n$ is of main interest; the number n is also essential information; σ assumed known.[1]
	(c). 2 sigma limits (d). 3 sigma limits	Commonly used bounds of imprecisions; usually used when σ known, or when n large.
	(e). Confidence interval (indicate one- or two-sided)	Data points assumed to be normally distributed; report confidence coefficient (level) $100(1-\alpha)\%$.[2]
	(f). Half-width of confidence interval (or confidence limits)	Same as (e) above; for symmetrical two-sided intervals; an index to bounds of imprecision.[2]
	(g). Probable error of the reported value	Probable error $= .6745 \frac{\sigma}{\sqrt{n}}$ for normally distributed data points when σ known. Use of $\sigma/\sqrt{n}$ preferred. Incorrect if σ not known.
	(h). Mean deviation, or average deviation, of a measurement from the mean calculated from the sample	Limiting mean of mean deviation $= \sqrt{\frac{2}{\pi}} \sqrt{\frac{n-1}{n}}\,\sigma$ for normally distributed data points when σ known. Use of σ usually preferred.
	(i). Any of the above expressed in percent, or ppm of $\bar{x}_n$.	State what is being expressed in percent, eg., $(\sigma/\sqrt{n})(100/\bar{x}_n)$, $\bar{x}_n$ being a fairly constant value.
m means each computed from n measurements	(j). (b), (c), (d) and (f) above	If the measurements are of equal precision and σ unknown, use $s_p^2 = \frac{1}{m} \sum_{i=1}^{m} s_i^2$ as estimate of σ^2. The no. of degrees of freedom associated with s_p is m(n−1).
	(k). Sample coefficient of variation ($v = \frac{s}{\bar{x}_n}$) or relative percent (v x 100)	Appropriate when the m means cover a wide range and where the v's computed for the m sets are about the same magnitude. Give range of v's for the m sets. The means must be positive and bounded away from zero.
Weighted mean $\bar{\bar{x}} = \frac{w_1\bar{x}_1 + w_2\bar{x}_2}{w_1 + w_2}$	(l). Standard error ($\sigma_{\bar{\bar{x}}}^2$) of the weighted mean	If $w_1 = 1/\sigma_{\bar{x}_1}^2$ and $w_2 = 1/\sigma_{\bar{x}_2}^2$, then $\sigma_{\bar{\bar{x}}}^2 = \frac{1}{w_1 + w_2}$ Not recommended when the σ's are not known and are estimated by s computed from small number of measurements.
An equation (theoretical or empirical) fitted to data points by the method of least squares	(m). Standard deviation computed from the deviations (residuals) of data points from the fitted curve	Report n, the number of data points, and k, the number of constants fitted, $s^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2/(n-k)$, where $\hat{y}_i$ is the value on the fitted curve for the particular x_i.[3] Value of s usually given in computer print-out.
Constants (coefficients) in the equation fitted to the data points by the method of least squares	(n). Standard errors of the coefficients based on the standard deviation computed under (m)	Standard errors usually given in computer print-out. Report n and k as above.[3]

TABLE 1 – IMPRECISION STATEMENTS – (Continued)

Value reported	Index or Measure of Error	Remarks
A predicted point on the curve $\hat{y}$ for a particular x_0	(o). Standard error ($s_{\hat{y}}$) of the predicted point	For the straight line case, the computer print-out gives the variance-covariance matrix $\begin{pmatrix} s_{11} & \\ s_{12} & s_{22} \end{pmatrix}$. $s_{\hat{y}}^2 = s_{11} + 2s_{12}x_0 + s_{22}x_0^2$.[3] Report n and k.
A predicted observed value for a particular x_0	(p). Standard error of the predicted value of y	For the straight line case, $s_y^2 = s_{\hat{y}}^2 + s^2$ where $s_{\hat{y}}^2$ and s^2 are that given in (o) and (m) respectively.[3] Report n and k
Value of function of the arithmetic means of several measured variables	(q). Standard error calculated by the use of propagation of error formulas	Appropriate when errors of measurements are small compared to the values of variables measured. Use standard error of the means of the variables in the formulas.[4] Report number of measurements from which these standard errors are computed.
Percentage or proportion (r/n), r and n being counts	(r). Confidence limits of the true proportion P	Procedures for obtaining exact and approximate confidence limits are discussed in Chapter 7, NBS Handbook 91. State one-sided or two-sided.

TABLE 2 – SYSTEMATIC ERROR[5] (BIAS) STATEMENTS

Value reported	Index or Measure of Error	Remarks
Numerical value resulting from a measurement process	Reasonable bounds ascribed to the value originating from: (i). systematic error reliably established	Detailed discussions of systematic errors are always helpful. Positive wording is appropriate: ". . . is not in error by more than . . ." ". . . is accurate within ± . . ."
	(ii). systematic error estimated from experience or by judgment	Use modifier such as "believed", "estimated", "considered", to signify the conjectural nature of the statement.
	(iii). combination of a number of elemental systematic errors	State explicitly the method of combination such as "the simple sum of the bounds" or "the square root of the sum of squares".
	(iv). uncertainty in some fundamental constant	Give reference to the value of constant used.
	(v). uncertainty in calibrated values	Ascertain the meaning of the systematic and random components of the uncertainty from the calibration laboratory so that decisions on the uses of these components can be made from the correct interpretations.
	(vi). bias in the method of computation	Correct if feasible, or give the magnitude.

TABLE 3 – UNCERTAINTY STATEMENTS

Value reported	Index or Measure of error	Remarks
Numerical value resulting from a measurement process	Bounds to inaccuracy: (1). Systematic error and imprecision both negligible	Explicit expression of correctness to the last significant figure, interpreted as being accurate within ±½ units in the last significant figure given.
	(2). Imprecision negligible. Bounds on inaccuracy given to no more than two significant figures.	Sentence form preferred such as given under remark for (i) and (ii). Footnote needed if bounds are given in tabular form.
	(3). Systematic error negligible. Index of precision (b), (g), (h), (i), (k), or (n) stated to no more than two significant figures	State explicitly the index used and give essential information associated with the index. Qualify index calculated by the word "computed". Avoid using expressions of the form a ± b unless the meaning of b is explained fully immediately following or in footnote.
	(3'). Systematic error negligible. Bounds to imprecision (c), (d), (e), or (f) stated to no more than two significant figures.	Same as under (3).
	(4). Neither systematic error nor imprecision negligible. Two numerics indicating bounds to systematic error and index of imprecision respectively	(2) and (3) above separately stated.
	(4'). Bounds to systematic error and imprecision combined, indicating the likely inaccuracy of the value	(2) and (3') above where the two components either have been previously described, or explained immediately following (or in footnote).
	(5). Quoted from literature	State reference and give author's interpretation of the uncertainty; add remark if meaning unknown or ambiguous.

[1] If σ is not known, use the computed standard deviation s based on k measurements as an estimate of σ, where $s^2 = \frac{1}{k-1} \sum_{i=1}^{k} (x_i - \bar{x}_k)^2$. The number (k−1) is the degrees of freedom associated with s.

[2] For interpretation see Chapter 1, NBS Handbook 91, **Experimental Statistics**, by M. G. Natrella, 1963.

[3] For details see Chapter 5 (straight line), and Chapter 6 (multivariate and polynomial), NBS Handbook 91.

[4] For details see "Notes on the use of propagation of error formulas", by Harry H. Ku, NBS Journal of Research, Vol. 70C, No. 4, October-December, 1966.

[5] See "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems" by Churchill Eisenhart, NBS Journal of Research, Vol. 67C, No. 2, April-June, 1963, and "Systematic Errors in Physical Constants" by W. J. Youden, Physics Today 14, 1961.

FIGURE 1 – SUMMARY OF RECOMMENDATIONS ON EXPRESSIONS OF THE UNCERTAINTIES OF FINAL RESULTS

SYSTEMATIC ERROR AND IMPRECISION BOTH NEGLIGIBLE (CASE 1)

In this case, the reported result should be given correct to the number of significant figures consistent with the accuracy requirements of the situation, together with an explicit statement of its accuracy or correctness.

SYSTEMATIC ERROR NOT NEGLIGIBLE, IMPRECISION NEGLIGIBLE (CASE 2)

(a) Qualification of a reported result should be limited to a single quasi-absolute type of statement that places bounds on its inaccuracy;

(b) These bounds should be stated to no more than two significant figures;

(c) The reported result itself should be given (i.e., rounded) to the last place affected by the stated bounds, unless it is desired to indicate and preserve such relative accuracy or precision of a higher order that the result may possess for certain particular uses;

(d) Accuracy statements should be given in sentence form in all cases, except when a number of results of different accuracies are presented, e.g., in tabular arrangement. If it is necessary or desirable to indicate the respective accuracies of a number of results, the results should be given in the form a ± b (or $a\,{}^{+b}_{-c}$, if necessary) with an appropriate explanatory remark (as a footnote to the table, or incorporated in the accompanying test) to the effect that the ± b, or ${}^{+b}_{-c}$, signify bounds to the errors which the a's may be subject.

(e) The fact that the imprecision is negligible should be stated explicity.

NEITHER SYSTEMATIC ERROR NOR IMPRECISION NEGLIGIBLE (CASE 3)

(a) A reported result should be qualified by: (1) a quasi-absolute type of statement that places bounds on its systematic error; and, (2) a separate statement of its standard error or of an upper bound thereto, whenever a reliable determination of such value or bound is available — otherwise, a computed value of the standard error so designated should be given, together with a statement of a number of degrees of freedom on which it is based;

(b) The bounds to its systematic error and the measure of its imprecision should be stated to no more than two significant figures;

(c) The reported result itself should be stated, at most, to the last place affected by the finer of the two qualifying statements, unless it is desired to indicate and preserve such relative accuracy or precision of a higher order that the result may possess for certain particular uses;

(d) The qualification of a reported result, with respect to its imprecision and systematic error, should be given in sentence form, except when results of different precision or with different bounds to their systematic errors are presented in tabular arrangement. If it is necessary or desirable to indicate their respective imprecisions or bounds to their respective systematic errors, such information may be given in a parallel column or columns, with appropriate identification.

SYSTEMATIC ERROR NEGLIGIBLE, IMPRECISION NOT NEGLIGIBLE (CASE 4)

(a) Qualification of a reported value should be limited to a statement of its standard error or of an upper bound thereto, whenever a reliable determination of such value or bound is available. Otherwise, a computed value of the standard error so designated should be given, together with a statement of the number of degrees of freedom on which it is based;

(b) The standard error, or upper bound thereto, should be stated to not more than two significant figures;

(c) The reported result itself should be stated, at most, to the last place affected by the stated value or bound to its imprecision, unless it is desired to indicate and preserve such relative precision of a higher order that the result may possess for certain particular uses;

(d) The qualification of a reported result with respect to its imprecision should be given in sentence form, except when results of different precision are presented in tabular arrangement and it is necessary or desirable to indicate their respective imprecisions, in which event such information may be given in a parallel column or columns, with appropriate identification.

(e) The fact that the systematic error is negligible should be stated explicitly.

POSTSCRIPT

Over the intervening years since the publication of Eisenhart's and Ku's articles, it has become apparent that a few additional comments may be useful. It is equally apparent that a complete revision is neither necessary nor desirable inasmuch as the major thrust and content of the articles remain as valid and as appropriate as when first written. For this reason, these comments are made as a postscript.

Uncertainty Assessments Must Be Complete

The uncertainty of a reported value is meant to be a credible estimate of the likely limits to its actual *error,* i.e., the magnitude and sign of its deviation from the truth. As such, uncertainty statements must be based on as nearly complete an assessment as possible. This assessment process must consider every conceivable source of inaccuracy in the result.

A measurement process generally consists of a very complicated sequence of many individual unit operations or steps. Virtually every step in this sequence introduces a conceivable source of inaccuracy whose magnitude must be assessed. These sources include:

- Inherent stochastic variability of the measurement process;
- Uncertainties in standards and calibrated apparatus;
- Effects of environmental factors, such as variations in temperature, humidity, atmospheric pressure, and power supply voltage;
- Time-dependent instabilities due to gradual and subtle changes in standards or apparatus;
- Inability to realize physical model because of instrument limitations;
- Methodology procedural errors, such as incorrect logic, or misunderstanding what one is or should be doing;
- Uncertainties arising from interferences, impurities, inhomogeneity, inadequate resolution, incomplete discrimination, etc.;
- Metrologist errors, such as misreading of an instrument;
- Malfunctioning or damaged apparatus;
- Laboratory practice including handling techniques, cleanliness, etc.; and
- Computational uncertainties as well as errors in transcription of data, and other calculational or arithmetical mistakes.

This list should not be interpreted as exhaustive, but rather as illustrative of the most common generic sources of inaccuracy that may be present.

The various sources of inaccuracy are generally classified into sources of *imprecision* (random components) and sources of *bias* (fixed offsets). To which category a particular source should be properly assigned is often difficult and troublesome. In part, this is because many experimental procedures or individual steps in the overall measurement process embody both systematic and stochastic (random) elements. (For an alternative discussion that questions the need for a clear cut distinction between random and systematic components of uncertainty, see [7].) One practical approach is to classify the sources of inaccuracy according to how the uncertainty is estimated. In this way, sources of imprecision are considered to be those components which *can be* and *are* estimated by a statistical analysis of replicate determinations. For completeness, the *systematic uncertainty components* can be considered to be the residual set of conceivable sources of inaccuracy that are biased and not subject to random variability, and those that may be due to random causes but *cannot be* or *are not* assessed by statistical methods. The systematic category includes sources of inaccuracy other than biases in order to obtain a complete accounting of all sources of inaccuracy in the measurement process. Hence, it is meaningful to report a random uncertainty contribution, only if one has a computed statistic for the magnitude of its imprecision or random variation. Many sources of inaccuracy may exist consisting of several components from both the random and systematic categories and can be assessed only after consideration of the more fundamental processes involved. The uncertainty in the calibration of an instrument with a standard reference material, for example, would have not only components from the uncertainty in the standard itself, but also uncertainty components arising from the use of the standard in performing the calibration.

Assessment of Imprecision (Random Uncertainties)

Although the treatment and expressions of reporting the imprecision of measurement results were adequately covered in the original article, a number of points are of sufficient importance to deserve reemphasis.

The only way to assess realistically the overall imprecision is to make direct—or preferably, when possible, indirect—replicate determinations [1] and calculate an appropriate statistic such as the standard error of the mean. It is extremely important to be definite on what constitutes a "replicate determination" because the extent to which conditions are allowed to vary freely over successive "repetitions" of the measurement process determines the scope of the statistical inferences that may be drawn from measurements obtained [2, sec. 4.1]. When measurements of a particular quantity made on a single occasion exhibit closer mutual agreement than measurements made on different occasions so that differences between occasions are indicated, the value of the computed standard error of the mean of all the measurements obtained by lumping all of the measurements together will underestimate the actual standard error of the mean. A more realistic value is given by taking the arithmetic means of the measurements obtained on the respective occasions as the *replicate determinations* and calculating the standard error of their mean in the usual way [3, sec. 3.5].

In many situations, it may not be possible or feasible because of time and cost constraints to perform a sufficient number of completely independent determinations of the measurement result. For results derived from several component quantities, the individual imprecision estimates must be propagated to obtain the imprecision of the final result. It must be emphasized, however, that

these estimates of imprecision should not be based exclusively on the information derived from just the present measurements. Presently derived information should be added to the information accumulated in the past on the imprecision of the measurement process. In this way, more realistic and reliable canonical values of the imprecision statistics may be established over time. Ideally, every major step or component of the measurement process should be independently assessed. This would include not only the variability inherent in the particular measurement of concern, but also the imprecision arising from corrections, calibration factors, and any other quantities that make up the final result.

Assessment of Systematic Uncertainties

Although a general guideline for the approach to the assessment of systematic uncertainties can be formulated, there are, unfortunately, no rules to objectively assign a magnitude to them. For the most part, it is a subjective process. Their magnitudes should preferably be based on experimental verification, but may have to rely on the judgment and experience of the metrologist. In general, each systematic uncertainty contribution is considered as a quasi-absolute upper bound, overall or maximum limit on its inaccuracy. Its magnitude is typically estimated in terms of an interval from plus to minus δ about the mean of the measurement result. By what method then should the magnitude of these maximum limits be assigned? It may be based on comparison to a standard, on experiments designed for the purpose [4], or on verification with two or more independent and reliable measurement methods. Additionally, the limits may be based on judgment, based on experience, based on intuition, or based on other measurements and data. Or the limits may include combinations of some or all of the above factors. Whenever possible, they should be empirically derived or verified. The reliability of the estimate of the systematic uncertainty will largely depend on the resourcefulness and ingenuity of the metrologist.

The Need for an Overall Uncertainty Statement

Without deprecating the perils of shorthand expressions, there is often a need for an overall uncertainty statement which combines the imprecision and systematic uncertainty components. Arguments that it is incorrect from a theoretical point of view to combine the individual components in any fashion are not always practical. First, an approach which retains all details is not amenable for large compilations of results from numerous sources. And second, this approach shifts the burden of evaluating the uncertainties to users. Many users need a single uncertainty value resulting from the combination of all sources of inaccuracy. These users believe, and rightly so, that this overall estimate of inaccuracy can be most appropriately made by the person responsible for the measurement result. It must be emphasized, however, that there is no one clearly superior appropriate method for reporting an overall uncertainty, and that the choice of method is somewhat arbitrary. Several methods are commonly employed [5,6].

One method is to add linearly all components of the systematic uncertainty and linearly add the total to the imprecision estimate. Since the individual systematic uncertainties (δ_j) are considered to be maximum limits, it logically should be added to an imprecision estimate at a similar confidence level. That is, for example, the overall uncertainty u may be given by

$$u = [t_\nu(\alpha)]s + \sum_{j=1}^{q} \delta_j$$

where s is the computed standard error based on ν degrees of freedom, $t_\nu(\alpha)$ is the Student-t value corresponding to a two-tail significance level of α=0.05, 0.01, or 0.001 (depending on the practice in the measurement field concerned), and δ_j is the magnitude of the estimated systematic uncertainty for each of the identified q systematic uncertainty components. This approach probably overestimates the inaccuracy, but can be considered as an estimate of the maximum possible limits. For example, if someone estimated that five contributions of about equal magnitude made up the total systematic error, that person would have to be very unlucky if all five were plus, or all five were minus. Yet, if there was one dominant contributor, it might be a very valid approximation.

Two other approaches have also been widely used. These methods add in quadrature all of the systematic uncertainty components, and either add the resulting quantity *linearly* to the standard error estimate,

$$s + \sqrt{\sum_{j=1}^{q} \delta_j^2} \quad ,$$

or add it *in quadrature* to the standard error estimate,

$$\sqrt{s^2 + \sum_{j=1}^{q} \delta_j^2}$$

These are frequently considered (erroneously) to correspond to a confidence level with P=68%.

In another method, often termed the PTB approach [6], the component systematic uncertainties are assumed to be independent and distributed such that all values within the estimated limits are equiprobable (rectangular or uniform distribution) [8]. With these assumptions, the rectangular systematic uncertainty distributions can be convoluted to obtain a combined probability distribution for which the variance may be computed. This may then be combined in quadrature with that for the random uncertainty. In its simplest form, the uncertainty components are combined to form an overall uncertainty by

$$u = k\sqrt{s^2 + (1/3)\sum_{j=1}^{q} \delta_j^2} \quad ,$$

where k is customarily taken as 2 or 3. The above simple form is not appropriate when one of the component δ_j's is much larger than the others; in such a case it will be more informative to keep that component separate from the others and add it linearly.

A Concluding Thought

If there is one fundamental proposition for the expression of uncertainties, it is

> The information content of the statement of uncertainty determines, to a large extent, the worth of the final result.

This information content can be maximized by following a few simple principles:

BE EXPLICIT

PROVIDE DETAILS

DON'T OVERSIMPLIFY

When an overall uncertainty is reported, one should explicitly state how the separate components were combined. In addition, for results of primary importance, a detailed discussion and complete specification of all of the separate uncertainty components is still required. In this way, some users will benefit from having the metrologist's estimate of the overall uncertainty, while more sophisticated users will still have access to all of the information necessary for them to evaluate, combine, or use the uncertainties as they see fit.

REFERENCES AND NOTES

[1] Youden, W. J. Statistical aspects of analytical determinations. The Analyst 77: 874–878; 1952, Dec: Reprinted in Journal of Quality Technology 4: 45–49; 1972, Jan: Youden, W. J., Connor, W. S., Making one measurement do the work of two; Chemical and Engineering Progress 49: 549–552; 1953, Oct. Reprinted in Journal of Quality Technology 4: 25–28; 1972, Jan.

[2] Eisenhart, Churchill. Realistic evaluation of the precision and accuracy of instrument calibration systems. J. Res. Nat. Bur. Stand. (U.S.) 67C (2): 161–187; 1963; Reprinted as paper 1.2 in NBS Special Publication 300-1.

[3] Eisenhart, Churchill. Contribution to panel discussion of adjustments of the fundamental physical constants. Langenberg, D. N., Taylor, B. N., eds. Precision Measurement and Fundamental Constants, Nat. Bur. Stand. (U.S.) Spec. Publ. 343: 509–518; 1971.

[4] Youden's burette experiment, Journal of Quality Technology 4: 20-23 1972, Jan. Youden, W. J., Systematic errors in physical constants. Physics Today, 14, 32–34, 36, 38, 40, 43; 1961, Sept. Reprinted as paper 1.4 in NBS Special Publication 300-1. Youden, W. J., Enduring values, Technometrics 14: 1–11; 1972, Feb.

[5] Campion, P. J.; Burns, J. E.; Williams, A. A code of practice for the detailed statement of accuracy. National Physical Laboratory. London: Her Majesty's Stationery Office. 1953. II-57.

[6] Wagner, Siegfried R. On the quantitative characterization of the uncertainty of experimental results in metrology: PTB-Mitteilungen 89: 83–89; 1979, Feb.

[7] Müller, Jörg G. Some second thoughts on error statements. Nuclear Instruments and Methods 163, 241–251; 1979.

[8] A numerical comparison of uncertainty limits resulting from these assumptions with those implied by several alternative distributional assumptions is provided by table 1 on page 184 of [2], and discussed on the same and following page.

Churchill Eisenhart
Ronald Collé
July 1980

NBS-114A (REV. 2-8C)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET *(See instructions)*	1. PUBLICATION OR REPORT NO.	2. Performing Organ. Report No.	3. Publication Date
	NBS SP 644		January 1983

4. TITLE AND SUBTITLE

Expression of the Uncertainties of Final Measurement Results: Reprints

5. AUTHOR(S)

Churchill Eisenhart, Harry H. Ku, and R. Collé

6. PERFORMING ORGANIZATION *(If joint or other than NBS, see instructions)*

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered

Final

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS *(Street, City, State, ZIP)*

Same as item 6.

10. SUPPLEMENTARY NOTES

Library of Congress Catalog Card Number: 82-600655

[ ] Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)*

This publication reprints and collects in one convenient source three articles, by NBS authors, that present a philosophical basis, general guidelines, and specific recommendations for expressing the uncertainties of final measurement results.

12. KEY WORDS *(Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)*

accuracy; errors; measurement uncertainty; precision; reporting of measurement data; systematic error; uncertainties.

13. AVAILABILITY

[X] Unlimited
[ ] For Official Distribution. Do Not Release to NTIS
[X] Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
[ ] Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES

19

15. Price

$3.00

USCOMM-DC 6043-P80

☆ U.S. GOVERNMENT PRINTING OFFICE : 1983 O—380-997 (2420)

NBS TECHNICAL PUBLICATIONS

PERIODICALS

JOURNAL OF RESEARCH—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. As a special service to subscribers each issue contains complete citations to all recent Bureau publications in both NBS and non-NBS media. Issued six times a year. Annual subscription: domestic $18; foreign $22.50. Single copy, $5.50 domestic; $6.90 foreign.

NONPERIODICALS

Monographs—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The principal publication outlet for the foregoing data is the Journal of Physical and Chemical Reference Data (JPCRD) published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

Consumer Information Series—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the **above** *NBS publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.*

Order the **following** *NBS publications—FIPS and NBSIR's—from the National Technical Information Service, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NBS Interagency Reports (NBSIR)—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161, in paper copy or microfiche form.

U.S. Department of Commerce
National Bureau of Standards

Washington, D.C. 20234
Official Business
Penalty for Private Use $300

U.S.MAIL

POSTAGE AND FEES PAID
U S DEPARTMENT OF COMMERCE
COM-215

FIRST CLASS